УДК 681.518

ОЦЕНКА КОЭФФИЦИЕНТА КОРРЕЛЯЦИИ НЕИСПРАВНОСТЕЙ ДИАГНОСТИРУЕМОГО ОБОРУДОВАНИЯ ЛОКОМОТИВА

М.С. Хазов

ВНИИ железнодорожного транспорта, г. Москва
E-mail: nikiforova@ocv.ru

Рассмотрены математические методы оценки корреляции неисправностей локомотивов. Представлен способ оптимизации множества диагностируемых параметров. Показаны возможности прогнозирования технического состояния диагностируемого объекта.

Существующие на сегодняшний день методы и средства технического диагностирования локомотивов позволяют оценить их техническое состояние после возвращения в депо из поездки. Однако значительный объём информации, получаемой в процессе движения, не используется. Источником подобной информации может служить разработанная бортовая система диагностирования. При оптимальном использовании дополнительной информации становится возможным повышение достоверности оценки технического состояния электровозов. Этого можно достичь путём увеличения количества обрабатываемых данных, установления зависимостей в причинах изменения технического состояния и на основе полученной информации прогнозировать возможные неисправности. На сегодняшний день актуальность поставленной задачи возрастает в связи с общим старением локомотивного парка, а также необходимостью повышения безопасности эксплуатации подвижного состава.

Для контроля работоспособности электроподвижного состава требуется подтверждение нормального функционирования всех его элементов. Решение этой проблемы усложняется необходимостью проведения значительных объёмов и количества контрольных операций [1]. Использование информации бортовой системы диагностирования, полученной в ходе проведения диагностических тестов электровоза, позволит перейти от плановых ремонтов к ремонтам по фактическому техническому состоянию, снизить затраты на техническое обслуживание и профилактику.

Решение поставленной задачи требует определения причин изменения технического состояния. Обычно для определения степени связанности двух событий используют значение коэффициента корреляции. Однако проблема, указывает ли полученная оценка коэффициента корреляции на зависимость двух событий с требуемой достоверностью или нет, остаётся и требует решения.

Сформулируем задачу следующим образом:

$$Y(k) = KX(d),$$

где Y – множество оценок технического состояния объектов контроля k, X – множество диагностируемых параметров d.

Требуется найти оператор отображения K множества X на множество Y. Ставится задача найти такой оператор K, при котором скалярное произведение множества Y и Y^*, должно стремиться к 1. Здесь Y^* действительное значение оценок технического состояния. Элементы оператора K являются коэффициентами корреляции между элементами множеств X и Y.

Корреляционный момент двух случайных величин x, y определяется по следующей формуле:

$$K_{xy} = \int_{-\infty}^{\infty} (x - m_x) f_1(x)dx \int_{-\infty}^{\infty} (y - m_y) f_2(y)dy.$$

где m_x, m_y – моменты выборок x и y, соответственно; $f_1(x)$, $f_2(y)$ – функции плотности распределения вероятности х и у, соответственно.

Пусть $(x_1,y_1),(x_2,y_2),...,(x_n,y_n)$ – выборка из n наблюдений пары случайных переменных (X, Y), которые имеют совместное двумерное нормальное распределение с коэффициентом корреляции q. Обозначим $\overline{x} = \Sigma x_j/n$, $\overline{y} = \Sigma y_j/n$. Коэффициент корреляции выборки r определяется как

$$r_{xy} = A\big/\sqrt{(BC)}, \qquad (1)$$

где

$$A = \sum_1^n \{(x_i - \overline{x})(y_i - \overline{y})\} = \sum_1^n x_i y_i - n\overline{x}\overline{y},$$

$$B = \sum_1^n (x_i - \overline{x})^2 = \sum_1^n x_i^2 - n\overline{x^2},$$

$$C = \sum_1^n (y_i - \overline{y})^2 = \sum_1^n y_i^2 - n\overline{y^2}.$$

Выборочное распределение коэффициента корреляции выборки имеет плотность распределения вероятности в точке r, заданную в виде

$$f_n(r,q) = \frac{2^{n-3}}{\pi(n-3)!}(1-q^2)^{n-1/2}(1-r^2)^{(n-4)/2} a(n,r,q),$$
$$-1 < r < 1,$$

где $$a(n,r,q) = \sum_{s=0}^{\infty} \Gamma^2\{(n+s-1)/2\}(2qr)^s\big/s!,$$

$\Gamma\{\}$ – гамма-функция, n – число наблюдений, q – коэффициент корреляции двух переменных (X,Y), r – коэффициент корреляции выборки (1).

Для характеристики связи между величинами (X, Y) в чистом виде перейдём от момента K_{xy} к безразмерной характеристике коэффициента корреляции

$$r_{xy} = \frac{K_{xy}}{\sigma_x \sigma_y},$$

где σ_x, σ_y – средние квадратичные отклонения величин X, Y.

Теперь следует определить оптимальное множество диагностируемых параметров, на котором достоверность оценки технического состояния достигает максимума. Оптимизация этого множества будет производиться путём исключения из него несвязанных данных, критерием отбора будет служить некоторое условие связанности данных.

Выдвинем гипотезу H о том, что полученный коэффициент корреляции действительно позволяет судить о связи имеющихся данных. Когда необходимо проверить, указывает ли наблюдаемое значение r на действительную коррелированность данных, то соответствующая нулевая гипотеза H имеет вид:

$$H: \ q=0.$$

Достаточно большое по абсолютной величине r будет стремиться опровергнуть нулевую гипотезу [2].

В рассматриваемой задаче обнаружения зависимости между различными типами неисправностей и измерений опровержение этой гипотезы приведет к подтверждению существования указанной зависимости. На этот вопрос можно ответить с помощью преобразования

$$t = \frac{r_{xy}}{\sqrt{1 - r_{xy}^2}} \sqrt{(n-2)}. \qquad (2)$$

Итак, если корреляция двух каких-либо событий превышает допустимый порог, то полагаем их взаимосвязанными, т.е., например, возникновение неисправности A влечет за собой появление неисправности B, или же обе эти неисправности являются следствием, ранее произошедшего события C. Это приведёт в перспективе к более точному прогнозированию событий и выявлению связей между возникновением неисправностей в одном агрегате и сбоев в работе другого. Что в свою очередь увеличит эффективность диагностирования.

При нулевой гипотезе выборочное распределение статистики неисправностей есть распределение Стьюдента с n–2 степенями свободы. Большие абсолютные значения r отвечают большим абсолютным значениям t, а поскольку выборочное распределение t симметрично относительно точки 0, то становится возможным определить уровень значимости SL:

$$SL=P(T\geq|t|)+P(T\leq-|t|)=2P(T\geq|t|),$$

где T подчиняется распределению Стьюдента с n–2 степенями свободы, а t выбирается в соответствии с (2) по выборочному коэффициенту корреляции r [2].

Таким образом, был получен уровень зависимости между событиями, который позволит получить однозначный ответ на вопрос, что может служить причиной отказа. В контексте задачи диагностирования ими могут быть: неисправности, включая время и место их возникновения, наличие внешних воздействий и т.д.

Для решения задачи оптимизации множества диагностируемых параметров следует исключить все данные, значения коэффициентов корреляции которых ниже уровня значимости. Это позволит избавиться от лишнего анализа несвязанных событий. Оптимизация представляет собой итерационный процесс, на каждом шаге которого происходит сравнение и оценивание параметров.

Пусть, например, X – множество отказов и измерений, а Y множество отказов, т.е. Y – подмножество X. Найдём такие X^* – исключаемые элементы множества X,

$$x_i^* \in X^*, \quad i = \overline{1, N},$$

для которых $r_{ij}<SL_{ij}$.

Исключив их из рассматриваемого множества, получим новые оптимизированные множества X_0 и Y_0.

В качестве примера возьмём выборочный коэффициент кор реляции между двумя оцениваемыми параметрами r=–0,63. Соответствующее значение t, вычисленное по формуле (2) для набора из 20 наблюдений равно –3,43. Уровень значимости составляет $2P(T_{20}\geq5{,}44)$. Значению t=–3,43 в соответствии с распределением Стьюдента отвечает значение SL, меньшее 0,01, что значимо. Нулевая гипотеза отклонена, а существование корреляции установлено.

Применение соответствующих математических методов даёт возможность прогнозировать техническое состояние диагностируемого объекта и позволяет сократить множество элементов объекта диагностирования и, как следствие, повысить точность в определении неисправностей, что приведёт к повышению безопасности движения.

Заключение

На практике учёт коэффициента корреляции неисправностей диагностируемого оборудования позволит повысить достоверность диагностической информации, улучшить качество диагностирования агрегатов и узлов, выявить связи между различными событиями. Это позволит продлить срок работы локомотива между ремонтами, снизить степень износа оборудования, экономить материальные и трудовые ресурсы.

СПИСОК ЛИТЕРАТУРЫ

1. Мозгалевский А.В., Гаскаров Д.В. Техническая диагностика. – М.: Высшая школа, 1975. – 248 с.
2. Справочник по прикладной статистике / Под ред. Э. Ллойда, У. Ледермана. – М.: Финансы и статистика, 1990. – 510 с.